THE RELATION BETWEEN BIRTH RATE AND DEATH RATE IN A NORMAL POPULATION AND THE RATIONAL BASIS OF AN EMPIRICAL FORMULA FOR THE MEAN LENGTH OF LIFE* GIVEN BY WILLIAM FARR.

By Alfred J. Lotka.

Both birth rate and death rate are, for obvious reasons, functions of the age-distribution in a given population.

A certain interest attaches to this circumstance from several points of view.

It is of interest to the statistician who seeks to "correct" the so-called "crude" death rates so as to bring them to a definite and common standard for purposes of comparison.†

Secondly, the subject is of interest in connection with certain problems of biochemistry, the growth of certain colonies of living organisms having been found to take place in accordance with the law of monomolecular reactions. As has been pointed out by Eijkman,‡ the question there arises, how this law of growth can be harmonized with the influence of the age-distribution on birth rates and death rates.§

Thirdly, the relation between age-distribution and birth and death rates is of interest in connection with the mathematical theory of evolution. For a quantitative measure of the "fitness" of a given type of organism, or of its "adaptation to given conditions," is given in its rate of growth under those conditions.|| But this rate of growth depends, not only on the fundamental characteristics of the species, but also on a factor

* Expectation of Life at birth.

† Cf. G. H. Knibbs, The Establishment of a Series of Norms for Death Rates. Trans. XV. Internat. Congr. Hyg. and Demogr. (1912), v. VI, p. 343; Knibbs & Wickens, The Establishment and Use of Population Norms, etc. *Loc. cit.* p. 352.

A. J. Lotka. A Natural Population Norm. Jl. Washington Ac. Sc. 1913, v. 3, pp. 241, 289.

‡ Cf. C. Eijkman, On the Reaction of Velocity of Organisms. Proc. Amsterdam Ac. Sci., 1912, p. 629; Reichenbach. Zeitschr, f. Hyg. u. Inf. Krankh, 1911, v. 69, p. 171; Wo. Ostwald, Uber die Zeitlichen Eigenschaften der Entwickelungsvorgänge, Leipsic, 1908; T. B. Robertson, On the Normal Rate of Growth of an Individual, Arch. f. Entwickelungsmechanik 1908, p. 108.

§ In ordinary chemical reactions and radioactive transformations this question does not arise, as the "force of mortality" of a species of molecules is independent of the age of the molecules. See A. J. Lotka, Am. Jl. Sci., 1907, v. 24, p. 205.

|| Cf. A. J. Lotka, Am. Jl. Science, 1907, v. 24, p. 216; Ann. Nat. Phil., 1911, p. 59; Jl. Wash. Ac. Sci., 1912, pp. 2, 49, 66; 1915, pp. 360, 397.

which, from this point of view, must be regarded as purely adventitious, namely the age-distribution among the individuals of the species.

The problem here arises of eliminating this adventitious factor from the case under consideration.

Fortunately it very largely eliminates itself by a spontaneous process: It is evident that the age distribution in a population is not fortuitous. It is capable of variation and fluctuation, but not without all rule and beyond all bounds. In point of fact, in an isolated population, where emigration and immigration are absent or negligible, the age-distribution tends to approach a certain definite normal type,* and when this is attained a relation between birth rate and death rate exists which no longer explicitly involves the age-distribution. Such a population behaves as if its birth rate and death rate were independent of the age-distribution, and grows and diminishes according to the exponential law or the law of geometric progression,† the same law which governs the progress of a monomolecular chemical reaction. There would therefore seem to be no occasion for surprise at the observation discussed by Eijkman, *loc. cit.*, that the growth of certain organisms follows this law irrespectively of the age-distribution in the colony.

While in practice the "normal" age-distribution may be more or less closely approached, it will, for obvious reasons,

* Cf. Poynting & Thomson, Text-book of Physics, Volume on Heat, 1904, p. 135; A. J. Lotka, Am. Jl. Sci., 1907, v. 24, p. 201; Science, 1907, v. 26, p. 21; Jl. Wash. Ac. Sci., 1913, pp. 241, 289; Sharpe & Lotka, Phil. Mag., April, 1911, p. 435.

† Cf. also M. Block, Traité de Statistique, 1886, p. 209; A. J. Lotka, Am. Jl. Sc., 1907, v. 24, p. 201.

It is evident that the birth rate and death rate can be *actually* independent of the age-distribution only if the force of mortality is independent of the age, *i. e.*, in the notation of equation (1) in the text, p. 123, if

$$p(a) = e^{-ka}$$
$$k = \text{const.}$$

When the age-distribution is fixed, and the population *behaves as if* its death rate were independent of the age-distribution, we can speak of an *equivalent constant force of mortality*, the value of which follows from equation (4), Am. Jl. Sc., 1907, v. 24, p. 200, viz:

$$\begin{aligned} d &= -\int_0^\infty c(a) \frac{d \log p(a)}{da} da \\ &= k \int_0^\infty c(a)\, da \\ &= k \end{aligned}$$

so that the equivalent constant force of mortality is numerically equal to the death rate per head.

never be exactly attained in any real population. But for the purposes of theoretical discussion, where we wish to eliminate from consideration the influence of the adventitious factor of age-distribution, it is convenient to base one's reflections on the example of an ideal population in which the "normal" age-distribution is established. This is all the more legitimate, since it has been shown that in practice this age-distribution may at times be very closely approached.*

A roughly approximate expression for the relation between the birth rate and death rate in a population which is increasing in geometrical progression (and in which, therefore, the age-distribution is normal) has been given by Bristowe.†
This expression is derived on the assumption that each individual attains the mean length of life, and is quite inexact.

The exact relation between the birth rate and death rate in a normal population has been given by the author ‡ in the form

$$\frac{1}{b} = \int_0^{\infty} e^{-ra} \quad p(a)da \tag{1}$$

where

b = birth rate per head per annum
r = natural rate of increase per head per annum
$= b - d$
d = death rate per head per annum
$p(a)$ = probability at birth of reaching age a (tabulated in the life tables and there commonly denoted by the symbol l_x).

Equation (1) expresses the relation between b and d in implicit form and is inconvenient for purposes of numerical computation.

In a later publication § another expression was therefore derived from (1) by expanding e^{-ra} under the integral sign and integrating term by term.

The formulae thus obtained give the relation between b and r and d and r in explicit form and offer certain advantages for purposes of computation, but are still somewhat compli-

* A. J. Lotka, Jl. Wash. Ac. Sc., 1913, pp. 243, 244.

† St. Thomas' Hosp. Report, 1876, quoted in Newsholme, Vital Statistics, 1899, p. 295.

‡ Am. Jl. Sci., *loc. cit.*

§ Jl. Wash. Ac. Sci., 1913, pp. 241, 289. Equations of the form (1) have also been dealt with by Pincherle, Mem. Bologna Ac. Sc., 1887, v. 8, p. 10; Bateman, Proc. Cambridge Phil. Soc., 1910, v. 15, p. 424.

cated, since the series converge, for common values of r, at such rate as to necessitate the retention of five or six terms, and the somewhat laborious determination of a corresponding number of constants.

A very much simpler form of the relation between b and d in a normal population, which holds with a very fair degree of approximation, has now been found, suggested by some reflections on an empirical (approximate) formula for the mean length of life quoted by Newsholme and attributed by him to Farr.* At the same time, following up this clue, an insight was gained into the *raison d'être* of Farr's empirical formula.

Farr showed that during the decade 1881–1890, the birth rate and death rate in England and Wales were related to the mean length of life† L very nearly according to the equation

$$\frac{1}{3}\cdot\frac{1}{b}+\frac{2}{3}\cdot\frac{1}{d}=L \tag{2}$$

This equation is of the form

$$\frac{P}{b}+\frac{Q}{d}=L \tag{3}$$

or

$$\frac{p}{b}+\frac{q}{d}=1 \tag{4}$$

By a simple transformation, we have, from (4)

$$(b-p)\ (d-q)=pq \tag{5}$$

which would indicate that for a given value of L the relation between b and d is, as a first approximation, hyperbolic in form.

Indeed, a glance at curve I in the accompanying figure, which is drawn in accordance with the data published by the writer in his paper on *A Natural Population Norm,* confirms this indication at least qualitatively. This curve shows the graph of the relation

* Newsholme, Vital Statistics, 1889, p. 301.

† For further discussion of the relation between the mean length of life and the rate of increase of a population see W. J. Spillman, Science, 1912; republished, with editorial note by the present writer, in Sci. Am. Suppl., 1912, p. 378.

$$\frac{1}{b}=\int_0^\infty e^{-(b-d)a} \quad p(a)da \tag{1a}$$

$$=\int_0^\infty e^{-ra} \quad p(a)da \tag{1}$$

for the decade 1871–1880 in England and Wales.

To obtain further light on this point, we proceed as follows: Equation (1a) is an implicit relation between b and d.

We may write it in the form:

$$\frac{1}{b}=F_1(r) \tag{6}$$

$$\frac{1}{d}=F_2(r) \tag{7}$$

whence, by eliminating r between (6) and (7) we should obtain

$$\frac{1}{b}=F_3\left(\frac{1}{d}\right) \tag{8}$$

Now F_1 and F_2 are known in the form of series, can, in fact, be obtained by expanding e^{-ra} under the integral sign in (1) and integrating term by term. The value of the coefficients of successive powers of r can be determined by mechanical quadrature when $p(a)$ is given in form of a life table and r is known.

We thus obtain, for England and Wales, 1881–1890

$$\frac{1}{b}=L+Ur+Vr^2+Wr^3+ \quad . \quad . \quad . \tag{9}$$

$$\frac{1}{d}=L+ur+vr^2+wr^3+ \quad . \quad . \quad . \tag{10}$$

Neglecting terms in r^2 and higher powers of r, and eliminating r between (9) and (10) we have:

$$\frac{u}{b}-\frac{U}{d}=L(u-U) \tag{11}$$

$$\frac{u}{u-U}\cdot\frac{1}{b}-\frac{U}{u-U}\cdot\frac{1}{d}=L \tag{12}$$

an equation which corresponds in form to Farr's formula (2) or its generalizations (3) and (4). This correspondence extends not only to the general linear form but also to the fact that the sum of the coefficients of $\frac{1}{b}$ and $\frac{1}{d}$ in (12) is unity, just as in Farr's formula (2).

NUMERICAL DATA.

1. Farr applies his formula to the period 1881–1890 (England and Wales).

The writer has determined by the method indicated in his paper cited above the coefficients u, U. There is first obtained

$$b = .02200 + .7130r + 6.766r^2 - 17.47r^3 + \quad . \quad . \quad . \tag{13}$$
$$d = .02200 - .2870r + 6.766r^2 - 17.47r^3 + \quad . \quad . \quad . \tag{14}$$

This gives, neglecting the terms in r^2 and r^3,

$$\frac{1}{d} = 45.45 + 593.0r + \quad . \quad . \quad . \tag{15}$$

$$\frac{1}{b} = 45.45 - 1473r + \quad . \quad . \quad . \tag{16}$$

$$.713\,\frac{1}{d} + .287\,\frac{1}{b} = 45.45 \tag{17}$$

so that we have for the mean length of life

$$L = .713\,\frac{1}{d} + .287\,\frac{1}{b} \tag{18}$$

The observed mean values of b and d during this period were

$$\left.\begin{array}{l} b = .03234 \\ d = .0190 \end{array}\right\} \tag{19}$$

Substituting these in (18) gives

$$L = 37.5 + 8.8 = 46.3$$

Observed	$L = 45.5$
Discrepancy	.8

Farr's formula gives

$$L = 45.6$$

an exceedingly close fit, which, however, is undoubtedly accidental.

2. For the period 1871–1880 the author's previous work furnishes detailed data for the male population.

It appeared of interest to recompute the graph of the relation between b and d by the hyperbolic formula (12) and compare it with the graph of the more exact relation (1) previously plotted.

The numerical data in this case are as follows:

$$b = .02418 + .7673r + 6.823r^2 + \quad . \quad . \quad . \tag{20}$$
$$d = .02418 - .2327r + 6.823r^2 + \quad . \quad . \quad . \tag{21}$$

This gives

$$.767\,\frac{1}{d} + .233\,\frac{1}{b} = 41.35 \tag{22}$$

This, written in hyperbolic form, gives

$$(b - .00563)\ (d - .01855) = .00010443 \tag{23}$$

or, putting

$$B = b - .00563 \tag{24}$$
$$D = d - .01855 \tag{25}$$

we have

$$BD = .00010443 \tag{26}$$

Table I below shows, for successive arbitrarily chosen values of b, the corresponding values of d calculated by (23), and, for comparison, also the corresponding values of d calculated by (1). For comparison also there are shown in the last two columns the corresponding products BD. The hyperbolic formula, of course, makes this product constant. The series formula gives a certain departure from constancy, which on inspection was found to be approximately equal to $-2r^3$.

This suggested a simple empirical formula which gives a very close approximation to the theoretically exact relation (1), namely

$$(b - .00563)\ (d - .01855) = .00010443 + 2r^3 \tag{27}$$

or

$$(b - .00563)\ (b - r - .01855) = .00010443 + 2r^3 \tag{28}$$

a relation which enables us to compute a series of values of b corresponding to given values of r. The corresponding values of d then follow directly in view of the relation $(b - d) = r$.

TABLE I.

I b	II d exact	III d hyperb.	IV* $BD\times 10^8$ exact	V* $BD\times 10^8$ hyperb.
.040	.02197	.02158	11754	10443
38	2205	2178	11329	10443
36	2217	2199	10994	10443
34	2234	2223	10752	10443
32	2256	2251	10574	10443
30	2284	2284	10454	10443
28	2322	2322	10447	10443
26	2368	2368	10450	10443
24	2423	2424	10434	10443
20	2574	2582	10332	10443
16	2796	2862	9758	10443
12	3125	3495	8090	10443
.008	3603	.06262	4143	10443

* $\begin{cases} B=b-.00563 \\ D=d-.01855 \end{cases}$

TABLE II.

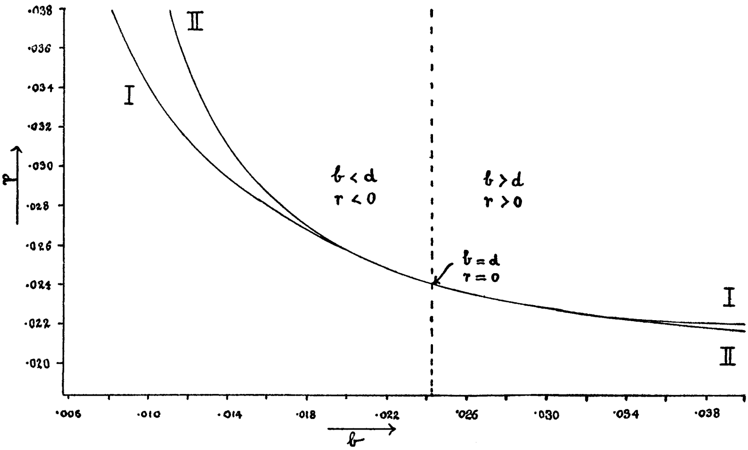
I r	II b exact	III b corr. hyp.	IV d exact	V d corr. hyp.	VI $BD\times 10^8$ exact	VII $BD\times 10^8$ corr. hyp. $=10443+2r^3\times 10^8$
.01803	.040	.04000	.02197	.02197	11754	11615
1595	38	3798	2205	2203	11329	11258
1383	36	3599	2217	2216	10994	10972
1166	34	3400	2234	2234	10752	10760
0944	32	3201	2256	2257	10574	10609
0716	30	3002	2284	2286	10454	10516
0478	28	2801	2322	2323	10447	10465
0232	26	2600	2368	2368	10450	10446
−.00023	24	2400	2423	2423	10434	10443
− 574	20	2003	2574	2577	10332	10405
− 1196	16	1617	2796	2813	9758	10101
− 1975	12	1225	3175	3200	8408	8902
−.02803	.008	.00799	3603	3602	4143	6039

Table II above shows, for the values of r corresponding according to equation (1) to successive entries in Table I, the values of b and d calculated both by the exact formula (1) and also by the corrected hyperbolic formula (28). The last two columns show the product BD as given by the exact formula (1) and also by the corrected hyperbolic formula (28). A comparison of columns II, III of Table I shows that, except for extremely small values of b, the hyperbolic formula gives very fair agreement with the exact formula.

On the other hand columns II, III and IV, V of Table II show an almost perfect agreement, which is somewhat remarkable in view of the empirical nature of the very simple correcting term $2r^3$.

The same features are brought out graphically in the accompanying figure which needs no further explanation beyond that furnished in the caption.

The agreement between columns II, III and IV, V of Table II is so close that there is no object in showing the corresponding curves. They would practically coincide within the errors of drawing.



CURVE I. GRAPH OF $b=\frac{1}{\int_0^\infty e^{-(b-d(a}p(a)da}$

CURVE II. GRAPH OF $.767\,\frac{1}{b}+.233\,\frac{1}{d}=41.35=L$

or $(b-.00563)\ (d-.01855)=.00010443$

Note that within the region of commonly occnrring valnes of b and d the two curves are nearly coincident.

To the left of the dotted vertical line the death rate exceeds the birth rate, so that the population is diminishlng. The marked divergence of curves I and II on the extreme left is of no practical significance, since such rapid rates of decrease as here represented can not long continue, if they occur at all, in any real population.

Summary:

The implicit relation between birth rate per head b and death rate per head d

$$\frac{1}{b}=\int_0^\infty e^{-(b-d)a}\quad p(a)da \tag{I}$$

which was developed in a previous paper, is now shown to be equivalent to the explicit hyperbolic relation

$$(b-p)\ (d-q)=pq=\text{const} \tag{II}$$

in the neighborhood of $b-d=0$ and for values of b and d such as occur in practice.

The hyperbolic relation (II), expressed as a relation between $\frac{1}{b}$ and $\frac{1}{d}$, assumes the linear form

$$\frac{u}{u-U}\cdot\frac{1}{b}-\frac{U}{u-U}\cdot\frac{1}{d}=L \tag{III}$$

which is of the same form as Farr's empirical rule for the mean length of life L, viz.:

$$\frac{1}{3}\cdot\frac{1}{b}+\frac{2}{3}\cdot\frac{1}{d}=L \tag{IV}$$

The empirical constants $\frac{1}{3}$ and $\frac{2}{3}$ which occur in Farr's rule thus receive a rational interpretation.